

Comments are welcome. This is for proceedings or a paper later on, so any comments are very welcome.

The two Figures are at:

Fig.1 (time series) at

<ftp://ftpprd.ncep.noaa.gov/pub/cpc/wd51hd/miscel/heightbreedtimeseries.gif>

Fig.2 (Maps of modes 1 and 2) at

<ftp://ftpprd.ncep.noaa.gov/pub/cpc/wd51hd/miscel/breed1.gif>

You could look at modes 3-6 if you wish at.

<ftp://ftpprd.ncep.noaa.gov/pub/cpc/wd51hd/miscel/breed2.gif>

<ftp://ftpprd.ncep.noaa.gov/pub/cpc/wd51hd/miscel/breed3.gif>

The leading EOFs and EOTs alluded to are to be seen at

<ftp://ftpprd.ncep.noaa.gov/pub/cpc/wd51hd/ake/daily/normeofl.gif>

and

<ftp://ftpprd.ncep.noaa.gov/pub/cpc/wd51hd/ake/daily/norm1.gif>

Calculating the Fastest Growing Modes by Empirical Means.

H. M. van den Dool

1. Introduction

Traditionally instability of atmospheric flow has been gauged by supplying a particular perturbation to a linearized dynamical operator. The operator is based either on a simplified analytical version of the governing equations, or on a numerical model and a given basic state (usually assumed constant). Here we will explore the appearance of the fastest growing (or least damped) modes in an operator based on data. It turns out that familiar low frequency modes, such as PNA and (N)AO like structures, can easily be cultured from daily data as complex modes with an overall growth rate, a period, two spatial maps, and two associated time series. It is thus suggested that these structures are (almost) unstable modes that grow by drawing energy from the mean flow (a full 3D basic state). As a slight departure from traditional studies we also argue that (1) the time series of the modes, although periodic, do not have to be sine and cosine and (2) the notion of Explained Variance in observed data by each mode separately does apply under certain restrictions.

2. The data.

Given is a space time data set $X(s,t)$, in this case X is the instantaneous daily 500 mb geopotential height taken from NCEP-NCAR Reanalysis - s is a spatial coordinate (5° lat by 10° lon grid), and t is time. We form anomalies by subtracting a harmonically smoothed 1979-1995 daily climatology, appropriate for the time of day, produced by Schemm et al (1997). We consider the domain 20N to the North Pole. The data set X is pruned as follows. In 1968 we take the field for January, 1, 3, 5, ... 23 at 0Z, i.e. twelve fields in one year. Similarly for 1969 through 1992, for a total of 300 fields during 25 years. We now have $X(s,t)$, where $t=1, 300$, representing a great diversity of NH January flow. The time t is a counter for both regular time and annual increments, all in one. (The calculations below are actually insensitive to scrambling t). $X(s,t)$ is considered the historical library. Formed in similar fashion is a 2nd distinct data set $X(s,t+\Delta t)$, which is the state of field X at a time Δt later - this is where knowledge about time evolution comes in. The time increment Δt is arbitrary, and while we have results for Δt ranging from 6 hours to 5 days, most results shown below are based on $\Delta t=2$ days.

3. Constructed Analogue

Given an Initial Condition, $X^{IC}(s, t_0)$ at time t_0 . We express $X^{IC}(s, t_0)$ as a linear combination of all fields in the historical library, i.e.

$$X^{IC}(s, t_0) \approx X^{CA}(s) \equiv \sum_{t=1}^{300} \alpha(t) X(s,t) \quad (1)$$

The determination of the weights $\alpha(t)$ for a non-orthogonal base $X(s,t)$ is non-trivial - this aspect is key to the construction of an analogue, the details can be found in Van den Dool(1994), see Appendix for a brief recipe of finding the weights. Except for some pathological cases, a set of 300 weights $\alpha(t)$ can always be found so as to satisfy the left hand side of (1) to within a tolerance

ϵ , which can be made as small as one wants. In order to find $\alpha(t)$ we need to truncate $X(s,t)$ in a reduced degree of freedom space and here we have chosen 50 EOTs (Van den Dool et al 2000), 50 such functions explain about 87% of the variance in the 0Z 1968-1992 data set. (50 EOFs explain 93.5% of the variance.)

4. Constructed Analogue Forecast

Equation (1) is purely diagnostic. We now submit that given the initial condition we can make a forecast with some skill by

$$X^F(s, t_0 + \Delta t) = \sum_{t=1}^{300} \alpha(t) X(s, t + \Delta t) \quad (2)$$

The calculation for (2) is trivial, the underlying assumptions are not. We ‘persist’ the weights $\alpha(t)$ resulting from (1) and linearly combine the $X(s, t + \Delta t)$ so as to arrive at a forecast to which $X^{IC}(s, t_0)$ will evolve over Δt . Note that the calculation of $\alpha(t)$ had nothing to do with Δt , so the forecast method is intuitive, and not based on minimizing some rms error for lead Δt forecasts. This is important in considering growing modes. While all statistical forecast methods, if based on minimizing rms error, damp the forecast anomaly amplitude to zero as skill goes down, the constructed analogue forecast is unconstrained and produces forecasts with amplitude out to very long lead.

5. Culturing growing modes

The purpose of this paper is not at all to dwell on the skill of this forecast method, but rather on the question as to which structure(s) appear when applying the construction method repeatedly. This repetition starts by making a new constructed analogue, this time to $X^F(s, t_0 + \Delta t)$ - this yields a new set $\alpha(t)$, which allows us to make a new forecast using (2). In short, writing F as short hand for $X^F(s, t_0 + \Delta t)$, one can write

$$F_{i+1} = \mathcal{P} \{ F_i \} \quad (3),$$

where forecast # $i+1$ is obtained from forecast # i by applying an operator \mathcal{P} , which combines step (1) and (2) into one. Along with F_i we have $\alpha(t)$ changing for each i . This process is stable if we renormalize at each iteration, i.e. make $|F_{i+1}| = |F_i|$ (where $||$ is a norm based on summing the square of F with cosine weighting over space). This avoids growth to infinity or damping to zero. In the process of expressing F_i at each iteration as a linear combination of $X(s,t)$ we freeze the annual cycle in January, or reset the clock in that aspect. After many iterations (hundreds or thousands, depending...), we save a sequence F_i , where $i = 501$ to 1000 for example, and analyze this synthetic data set by using a traditional EOT analysis. The counter i (iteration) may also be interpreted as time (but now in perpetual season mode), so we analyze a synthetic data set $F(s,t)$. In nearly all cases studied we appear to have converged to a single complex mode $M(s,t)$ which can be described as

$$M(s,t) = G [A(s) f(t) + B(s) g(t)]$$

where G is the overall growth, A and $B(s)$ are two spatial maps, and $f(t)$ and $g(t)$ are periodic time series multiplying each map. Here t equals the counter i . Growth G may or may not be exponential (e^{ot}) as in Simmons et al(1983) and Anderson(1991) - below G is simply expressed as percentage amplitude growth per 24 hours. Both the maps and the time series are to be

determined by the process described above. Compared to Simmons et al (1983), or Linear Inverse Modelling (Penland and Magorian 1993) we appear to have one less constraint, because the time series are not assumed to be sine and cosine. The maps A and B turn out to be spatially orthogonal, while f and g are temporally uncorrelated periodic functions with period T. The instantaneous growth rate defined as $|F_{i+1}|/|F_i|$ before renormalization of $|F_{i+1}|$, while averaging out to the growth rate σ , can be an arbitrary function of time and is periodic with period T/2. The structure that survives the iterations (i.e. comes out first) is either the fastest growing or the least damped mode. The time series look like a deformed sine/cosine pair, i.e. one time series is saw-tooth like, while the other has large residence time at the extremes, and fast zero crossings.

6. Results.

An example for $\Delta t=2$ days may clarify the above. Shown in Fig.1 are the g and f time series (blue and red) of the 1st mode in the synthetic data, while Fig.2 (top row) shows the two spatial maps. Loosely speaking we go from map A to map B in a quarter period, then to the negative of map A, the negative of map B and return to map A after T days, where T may be non-integer. In this case, map A looks like the PNA, while map B is somewhat like (N)AO, but we stress that the modes thus produced are not the same as leading EOFs or one point teleconnection maps. The overall growth rate is massaged out of this display, but can be seen in Fig.1 (green curve). In the mean this least damped mode decays at 4.5% per day, but the actual damping rate varies between 3.68 and 5.16% per day over T/2. The period of this 1st mode is roughly 630 days, clearly a low frequency mode, which, rather strikingly, was distilled here from daily data arranged in 23 day sequences at two day interval from 25 different years. The notion of a period (as in the period of a periodic function) still holds here, although the time series are not single harmonics, but have a spectrum. In terms of a general purpose computer code we found it difficult to determine T from the synthetic data $F(s,t)$.

The repeated application of operator \mathcal{P} to an arbitrary initial condition, with renormalization at each step, is similar to a power method to find the (complex) eigenvector of the (non-symmetric) \mathcal{P} . Finding the first mode was described above. We find the 2nd mode by removing from $X(s,t)$ the projection of $X(s,t)$ onto the first mode, and recalculating \mathcal{P} etc, etc, and so on for mode 3, 4, 5 and 6. Table 1 gives pertinent information about the first 6 modes. The period and the growth rate are classical attributes of normal modes. Here we add Explained Variance (EV) (in the $X(s,t)$) data, as an interesting and new side issue. One would like to know whether the fastest (least damped) growing modes are a curiosity, or that they really mean something in the world around us. One way of expressing that is to calculate EV, and in a sense we are lucky that, by our construction, the real and imaginary parts are orthogonal, and that all maps of subsequent modes are orthogonal to the previous maps, such that a unique EV by mode does exist. Table 1 shows long periods, mildly negative growth rates, and fairly high EV. While the order in which modes are selected is determined by growth rate alone, they still, as it happens so, order approximately in terms of EV. One might interpret this as a sign that, given random forcing and non-linearity, the least damped modes have the highest probability to maintain amplitude, and be naturally selected to ‘explain’ variance in the real world. The precise damping rates (5% loss per day) should not be taken that literally - for $\Delta t=6$ hours growth is 6% positive for leading modes.

7. Discussion

-)We found the resulting modes to be independent of the initial condition (except polarity and phase).

-) Sometimes (often!) the period is so long that a very long synthetic data series would have to be produced to determine whether or not the period is less than infinity. For all intents and purposes we rounded off to zero frequency, or $T=\infty$, if more than 1000 days of integration would have been needed. In the case of zero frequency, the oscillation is stuck in one map of fixed polarity plus overall growth or decay. The 2nd mode in Fig.2 turns out to be zero frequency. If we had produced data for $i=501$ to 600 only, we might have concluded that the two maps now combined into one complex mode #1 (top Fig.2) are two zero frequency modes. Indeed, when choosing $\Delta t=1$ or 3 days, the PNA and NAO like modes may appear as zero frequency, or coupled to each other or to yet another mode, but in all cases the period is very long. With such low damping rates it takes only minimal forcing to make these modes persist for a long time.
-) Sometimes the procedure described above does not appear to converge, i.e. the synthetic data $F(s,t)$ contains more than a single complex mode, no matter how long we iterate. This could either be a failure of the iteration method in a case where the growth rates of 2 or more different modes are very very close, or perhaps we need to entertain the thought of generalized modes that consist of more than 2 maps and 2 time series.
-) One could imagine doing the same experiment with a (any) numerical model. Take an initial condition, and declare it to be a Jan 15 IC. Make a model forecast for Jan 16 ($\Delta t=1$ day). Take departure from Jan 16 climatology and add Jan 15 climatology back in, and make a new forecast for Jan 16. Etc. In fact, for small Δt , this is one way of checking whether the eigenvectors (calculated directly from the operator) are correct. It is difficult to calculate more than one mode this way. Because the climatology may not be a model solution the procedure may actually involve 2 model forecasts, the difference between which is re-scaled after each forward step.
-) The non-sinusoidal character of the periodic time series has been discussed before (Frederikson and Branstator 2001; FB) - the distinguishing feature for this being that the basic state is not an absolute constant (Mak personal communication). In FB the annual cycle in basic state was invoked to argue non-constancy of the basic state, but in the approach reported here even the basic state in a perpetual run would be not be a constant in the sense that energy goes from the basic state to the perturbation. Indeed our analysis of data from a perpetual barotropic model shows growing modes with non-sinusoidal time series.

Figure Legends:

Fig.1: Time series of fastest growing mode 1, for $\Delta t = 2$ days, for an arbitrary portion of the time series slightly in excess of one period. Time series are scaled to vary from -1 to +1. Blue and red are the time series multiplying the real and imaginary part. The green curve represents the instantaneous growth rate (% amplitude change per day).

Fig. 2: The spatial maps of mode 1 (top row), and mode 2 (bottom row). On the left the real part, and on the right the imaginary part. The second mode has zero frequency and its imaginary part is zero. Units are gpm/100, multiplied further by the inverse of re-scaling (close to unity usually) applied to the time series.

Table 1. The period (days), the growth rate (% per day) and Explained Variance (%) in the original (untruncated) data of fastest growing modes 1 to 6 in January, 0Z 500 mb height, 20-90NH, and $\Delta t = 2$ days.

	Period	growth rate	Explained Variance		Cumulative EV
<i>Units: (Days)</i>		<i>(% per day)</i>	<i>(%)</i>		<i>(%)</i>
Mode #			Real	/ Imaginary	
1	630	-4.5	7.7	9.1	16.8
2	∞	-7.8	4.5	—	21.3
3	58	-8.2	4.8	3.9	30.0
4	28	-9.7	2.7	4.5	37.2
5	58	-11.3	5.5	3.6	46.1
6	22	-15.2	2.8	2.2	51.1

References:

- Anderson, J.L., 1991: The robustness of barotropic unstable modes in a zonally varying atmosphere. *J. Atmos. Sci.*, 48, 2393-2410.
- Frederiksen, Jorgen S., and Grant Branstator, 2001: Seasonal and Intraseasonal Variability of Large-Scale Barotropic Modes. *Journal of the Atmospheric Sciences*: Vol. 58, No. 1, pp. 50–69.
- Penland, C. and Th. Magorian, 1993: Prediction of Niño 3 Sea Surface Temperatures Using Linear Inverse Modeling. *Journal of Climate*: Vol. 6, No. 6, pp. 1067–1076.
- Schemm, J., H. van den Dool, J. Huang and S. Saha: Construction of daily climatology based on the 17-year NCEP/NCAR Reanalysis. *Proceedings of 1st Reanalysis workshop*. WMO/TD-No. 876, WCRP-104, pp290-293, Silver Spring, MD, 27-31, Oct. 1997.
- Simmons, A., J. M. Wallace, and G. W. Branstator, 1983: Barotropic wave propagation and instability, and atmospheric teleconnection patterns. *J. Atmos. Sci.*, 40, 2383-2398.
- Van den Dool, H. M., 1994: Searching for analogues, how long must one wait? *Tellus*, 46A, 314-324.
- Van den Dool, H. M., S. Saha and A. Johansson, 2000: Empirical Orthogonal Teleconnections. *J. Climate*, 13, 1421-1435.

Appendix

The coefficient $\alpha(t)$ in expression (1) are determined as follows:

- 1) The data set $X(s,t)$ is truncated to N EOFs, where N is 50 for daily Z500. Below $X(s,t)$ denotes truncated fields.
- 2) The initial conditions is likewise truncated by projection onto the same 50 EOFs.
- 3) The ‘normal’ equation is set up for the problem:

$$A \alpha = b \quad (A1)$$

where A is a 300 by 300 matrix with elements $a(i,j) = \sum X(s, t_i)X(s, t_j)$, $i=1,300$; $j=1,300$ and summation is over space with cosine weighting. (The summation may alternatively be performed in EOF spectral space.) The vector b has elements $b(j) = \sum X^{lc}(s)X(s, t_j)$, and α is the vector of length 300 we want to solve for.

- 4) The matrix A is ‘ridged’ by adding a small positive constant to the main diagonal. This constant is chosen large enough to make the resulting α reasonably insensitive to further changes in the details of the calculations (such as truncation at 51/49 EOF, using EOT instead of EOF, perturbing one of the $X(s,t)$, etc etc).
- 5) A1 is solved by standard linear algebra techniques.